

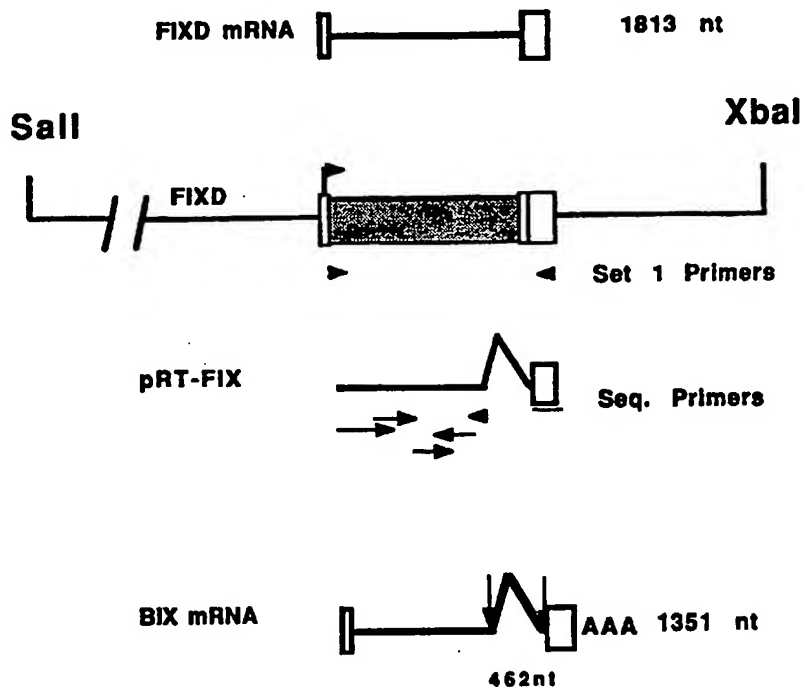


## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C12N 15/12, 15/63, 5/10, A01K 67/027,</b> <b>C12N 15/11</b>	<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 95/30000</b> <b>(43) International Publication Date:</b> 9 November 1995 (09.11.95)
<b>(21) International Application Number:</b> PCT/GB95/00996 <b>(22) International Filing Date:</b> 2 May 1995 (02.05.95) <b>(30) Priority Data:</b> 9408717.8                      3 May 1994 (03.05.94)                      GB <b>(71) Applicant (for all designated States except US):</b> BIOTECH- NOLOGY AND BIOLOGICAL SCIENCES RESEARCH COUNCIL [GB/GB]; Central Office, Polaris House, North Star Avenue, Swindon SN2 1UH (GB). <b>(72) Inventor; and</b> <b>(75) Inventor/Applicant (for US only):</b> CLARK, Anthony, John [GB/GB]; 29 Broomieknowe, Lasswade, Midlothian EH18 1LN (GB). <b>(74) Agents:</b> SHEARD, Andrew, Gregory et al.; Kilburn & Strode, 30 John Street, London WC1N 2DD (GB).		<b>(81) Designated States:</b> AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, EE, ES, FI, GB, GE, HU, IS, JP, KE, KG, KP, KR, KZ, LK, LR, LT, LU, LV, MD, MG, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, TJ, TM, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD, SZ, UG).  <b>Published</b> <i>With international search report.</i> <i>Before the expiration of the time limit for amending the</i> <i>claims and to be republished in the event of the receipt of</i> <i>amendments.</i>

**(54) Title:** DNA SEQUENCES**(57) Abstract**

Poor expression yields of recombinant human factor IX are attributable to aberrant splicing in heterologous expression systems such as transgenic hosts. The aberrant splicing sites have been identified as (a) a donor site including mRNA nucleotide 1085; and (b) an acceptor site including mRNA nucleotide 1547; adopting the mRNA nucleotide numbering of Figure 2 of the drawings. Improved factor IX expression sequences have at least one of these sites engineered out, so as to prevent or reduce the effect of aberrant splicing and to increase yields. The improved DNA sequences may also be useful in gene therapy.



**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

DNA SEQUENCES

5 This invention relates to DNA sequences encoding human factor IX (FIX). Such sequences are useful in expression systems for factor IX, including transgenic animals, and also have potential in gene therapy.

10 It is difficult to achieve high expression yields of factor IX in heterologous, particularly transgenic, systems. For example, while the basic approach to  $\beta$ -lactoglobulin-driven transgenic expression of human factor IX in the milk of transgenic animals such as sheep (as disclosed in WO-A-8800239) does work, the yields obtained are low. There seem to have been two main reasons for this:

20 **Failure to express.** The use of factor IX cDNAs has generally proved a problem in terms of getting reasonable levels of the appropriate FIX transcript. This problem was partially solved by the transgene rescue approach (described in WO-A-9211358, "Increased Expression by a Second Transferred Sequence in Transgenic Organisms"). In this prior publication, cointegration of  $\beta$ -lactoglobulin (BLG) with the human factor IX-encoding construct FIXD led to the production of lines of mice expressing high levels of FIXD mRNA. The milk of these animals, however, contained very little FIX.

30 **Aberrant splicing.** Closer inspection of the FIXD mRNA transcripts in the BLG+FIXD mice showed that they were approximately 450 bp shorter than predicted. It was surmised that these are deleted internally most probably by an aberrant splice of

the mRNA (Clark et al., *Bio/Technology* 10 1450-1454 (1992)).

5 Splicing of human factor IX mRNA in liver cells has been discussed in *J. Biol. Chem.* 270, 5276-5281 (1994) (Kurachi et al). Here it is indicated that the presence of splicing signal sequences results in increased expression of factor IX since spliceosome complexes act to protect precursor mRNAs from random degradation before  
10 being transported out of the nucleus.

It has now been determined that aberrant splicing is indeed a cause of low factor IX yield in heterologous or transgenic expression systems. Furthermore, and most  
15 significantly, the location of cryptic splice sites in the human gene encoding factor IX has been identified. This discovery enables factor IX-encoding DNA sequences to be engineered to avoid the observed aberrant splicing.

20 According to a first aspect of the present invention, there is provided DNA having a sequence encoding a protein having human factor IX activity, wherein the DNA is modified to interfere with the functioning of at least one of the following cryptic splice sites:

- 25 (a) a donor site including mRNA nucleotide 1086;  
and  
(b) an acceptor site including mRNA nucleotide 1547;

30 adopting the mRNA nucleotide numbering of Figure 2 of the drawings.

DNA in accordance with the invention makes possible much higher levels of fIX expression than hitherto described by correcting an aberrant splice of fIX sequences.

A donor site in nuclear pre-mRNA (that is, the primary transcript of the gene which exists transiently in the nucleus before splicing to generate mRNA which is exported to the cytoplasm) contains the nucleotides GU, which after splicing become the 5' terminal nucleotides of the excised intron. An acceptor site in the nuclear pre-mRNA contains the nucleotides AG, which after splicing become the 3' terminal nucleotides of the excised intron. The nucleotide numbers given in the preceding paragraph are for the G residue of the (5') donor site and the G residue of the (3') acceptor site, respectively.

Preferred DNA in accordance with the invention encodes wild-type human factor IX. However, DNA encoding variants (particularly allelic variants from a consensus sequence), conservative mutations or other proteins is also within the scope of the invention, provided that the proteins are substantially homologous with human factor IX. "Substantial homology", as is well understood in the art, may be assessed either at the protein level or the nucleic acid level. For example, at the protein level, substantial homology may be said to be present if a candidate protein exhibits amino acid homology to human factor IX at a level of at least 40, 60, 80, 90, 95 or 99%, in increasing order of preference. At the nucleic acid level, substantial homology may be said to be present if a candidate DNA sequence exhibits DNA sequence homology to human factor IX at a level of at least 80, 90, 95 or 99%, in increasing order of preference.

It will be appreciated that the invention has application to a variety of DNA sequences encoding factor IX (or another protein having factor IX activity). In

particular, the invention is applicable to cDNA sequences, genomic sequences having a full complement of natural introns and "minigene" sequences, containing some but not all of the introns present in genomic DNA  
5 encoding factor IX.

There are a variety of ways in which DNA in accordance with the invention may be modified to interfere with the functioning of the cryptic donor/acceptor sites so as to  
10 prevent or at least significantly reduce aberrant splicing.

First, the intron/exon structure of the constructs could be changed, on the basis that additional introns 5' or 3' would "compete" with the cryptic splice in some way.  
15 However, this approach may be relatively complex and lead to only partial suppression of aberrant splicing.

Secondly, the cryptic donor site could be engineered out.  
20 Either the G or the U of the mRNA donor site could be replaced with another base, or both could be replaced, provided that a stop codon does not result from the change. This approach is technically simpler than the competitive intron approach described above, but  
25 necessitates a change in the amino acid sequence of factor IX, because the GU residues at the donor site form the first two nucleotides of a valine codon, and all valine codons begin GU. This may not be a disadvantage, and may actually be an advantage if a second or  
30 subsequent generation variant of factor IX is being engineered. However, it is not suitable if retention of the wild-type factor IX sequence, at least in the region of the donor site, is essential.

Thirdly, and in most instances preferably, the cryptic acceptor site can be engineered out. This site lies in the 3' untranslated region of factor IX DNA, and so there are no implications for the amino acid sequence. Either the A or the G of the mRNA acceptor site could be deleted or replaced with another base, or both could be deleted or replaced. In fact, in some of the simplest embodiments of the invention, deletion of the acceptor site just requires the production of a factor IX cDNA segment which is shortened at the 3' end (or, of course, a DNA other than a cDNA shortened correspondingly). In other embodiments, site-directed mutagenesis techniques may be used specifically to alter the acceptor site (or, of course, the donor site).

DNA in accordance with the invention is useful in systems for expressing factor IX (or like proteins).

According to a second aspect of the invention, there is provided an expression host comprising DNA in accordance with the first aspect of the invention operably linked to an expression control sequence. The expression control sequence will usually comprise a promoter, and other regulatory sequences may be present.

While the invention may be generally useful across various different cell types and cultured cells, it is with transgenic animal expression systems that the invention has particular application, because of the large yields that are in principle available from this technology. Therefore, the expression host is in certain favoured embodiments an animal, such as a mammal.

5 A preferred transgenic system for the production of heterologous proteins involves the use of transgenic placental non-human mammals, especially sheep and other dairy animals, which express a transgene in the mammary gland (of an adult female) under the control of a milk protein promoter, particularly that of the milk whey protein  $\beta$ -lactoglobulin, as disclosed in WO-A-8800239, WO-A-9005188 and WO-A-9211385.

10 However, the invention is not limited to the use of these preferred transgenic systems. It is expected that factor IX-encoding sequences will be used in gene therapy approaches for haemophilia, for example using retroviral vectors or direct transfection techniques into stem  
15 cells. The advantages of an improved FIX sequence which does not aberrantly splice are self evident.

20 Preferred features for each aspect of the invention are as for each other aspect, *mutatis mutandis*.

The invention will now be illustrated by the following examples. The examples refer to the drawings, in which:

25 FIGURE 1 refers to Example 1 and shows the scheme used to confirm the aberrant splicing of the FIXD construct;

30 FIGURE 2 also refers to Example 1 and is adapted from Anson et al., *The EMBO Journal* 3(5) 1053-1060 (1984) and shows the locations of the cryptic donor and acceptor sites in factor IX mRNA;

FIGURE 3 refers to Example 1 and shows in more detail how the donor and acceptor sites interact;



the figure also shows generalised consensus sequences for donor and acceptor sites;

5       FIGURE 4 shows the gross structure of the human factor IX gene, including the locations of the cryptic splice sites;

10       FIGURE 5 refers to Example 2 and shows a PCR-based scheme for distinguishing between unspliced and aberrantly spliced mRNA for different constructs and in different expression systems;

15       FIGURE 6 refers to Example 3 and shows the construction of a construct designated FIXD- $\Delta$ 3'splice;

20       FIGURE 7 refers to Example 4 and shows a Western Blotting analysis of milk from transgenic mice expressing high yields of human factor IX. Milk samples from two animals from line FIXD $\Delta$ 3'-splice (31.31.2 and 31.3) were electrophoresed under non-reducing conditions. Milk samples were diluted 1/200 and either 5  $\mu$ l or 10  $\mu$ l loaded. fIX, 10 ng fIX; CM, control milk; CM+fIX, control milk + 10 ng fIX; and

25

30       FIGURE 8 also refers to Example 4 and shows Northern blots of representative RNA samples from FIXD- $\Delta$ 3'splice mice probed with a factor IX-specific probe. Mammary gland RNAs from high and medium expressing BIX mice (BIX33.1 and BIX34.1) were compared to mammary gland samples from FIXD $\Delta$ 3'-splice transgenic mice (labelled BIX $\Delta$ 3'3.10->BIX $\Delta$ 3'44.2). Blots were probed with labelled insert

from p5G3'CVII a plasmid containing cDNA sequences human FIX and then reprobred with GAPDH to control for loading. The sizes of the transcripts are indicated. The FIXD $\Delta$ 3'-splice transcripts are evidently larger than those from the BIX mice.

**EXAMPLE 1 - Aberrant Splicing of Construct FIXD**

The aberrant splicing of the FIXD mRNA was confirmed by cloning these transcripts by RT-PCR from mammary gland RNA of one of the expressing lines of mice. FIXD is disclosed in Example 3 of WO-A-9005188 and Comparative Example 6 of WO-A-9211385 and comprises human factor IX (FIX) cDNA fused to  $\beta$ -lactoglobulin (BLG) 5' and 3' sequences (including exons 6 and 7); FIXD contains no naturally occurring introns. Primers (Set 1: Figure 1) specific to the 5' end of the FIX cDNA and 3' end of BLG were designed and constructed. The primers had the following sequences:

Set 1-5'FIX (code no. 292343): 5' CAC CAA GCT TCA TCA CCA TCT GCC 3' \*  
Set 1-3'BLG (code no. 290646): 5' GGG TGA CTG CAG TCC TGG TCC C 3'  
\*contains an introduced *HindIII* site to enable cloning.

These primers amplified the shorter FIXD transcript (named BIX) from the BLG+FIXD mice and this was cloned in plasmid vector pBLUESCRIPT as pRT-FIX, which was then sequenced. The sequence of pRT-FIX showed a 462 nt internal deletion in the FIX sequences. Thus instead of the 1813 nt size of predicted for FIXD mRNA the BIX transcripts were 1351 nucleotides (Figure 1).

The sequence of pRT-FIX, determined by the dideoxy method of Sanger, identified the precise location of the deletion observed in BIX mRNA. Inspection of the FIX cDNA sequence (Anson et al., *The EMBO Journal* 3(5) 1053-

1060 (1984)) and comparison to the 5' and 3' break points deduced from pRT-FIX showed that the deletion was almost certainly due to aberrant splicing. Thus the deletion comprises bp 1085-1547 inclusive (as numbered in the Anson paper and in Figure 2 of this specification). The most 5' sequence is 5'GUAAGUGG and the most 3' sequence is UUUCUCUUUACAG3' (Figure 3). These are very 'good' consensus sequences for the donor (5') and acceptor (3') sites of an intron. (The 5' and 3' ends of an intron must have GU and AG respectively: these are absolute requirements for splicing; the other bases here are also close to the consensus for the donor and acceptor sites.)

Note that the presence of donor and acceptor sites does not mean that a gene must be spliced in this way: from the sequence one cannot predict whether or not a splice will occur. Indeed in the natural factor IX gene these sites are present in the last exon (exon 8) separated by the same sequences that are in FIXD (Figure 4). Nevertheless these sites are not used in the normal expressing factor IX pre-mRNA in human liver. Thus, for some reason the FIX transcripts produced in the mammary gland use these cryptic splice sites, resulting in the production of the internally deleted BIX mRNA. This internally deleted mRNA cannot code for a functional FIX protein since it results in the removal of segment coding for the last 109 amino acids of FIX.

**EXAMPLE 2 - Aberrant Splicing Occurs with Other FIX Constructs**

The identification of the aberrant splicing of FIX cDNA sequences was made with mice expressing the FIXD construct (cointegrated with BLG). Transgenic sheep with FIX cDNA sequences had previously been made, but in these

sheep the fIX cDNA sequences were integrated into the first exon of the intact BLG gene, as a construct called FIXA (as described in Example 3 of WO-A-8800239). This construct also appeared to behave rather poorly and produced rather low levels of fIX in the milk. It was, therefore, also of interest to see whether this aberrant splice occurred in the mammary gland with this fIX construct. Mammary RNA samples from sheep carrying another relatively poorly expressing construct, JFIXA1 (identified as J FIX A 1 in Section E of Example 4 of WO-A-9005188), were also procured from transgenic sheep derived from a founder transgenic prepared as disclosed in WO-A-9005188. A set of PCR primers (Set 2: Figure 5) were designed which upon RT-PCR amplification of RNA would distinguish the unspliced fIX sequences from the aberrantly spliced mRNA that was observed for BIX mRNA. In wild type (non-aberrantly spliced mRNA) these primers would generate a 689 p fragment, whereas in aberrantly spliced mRNA they would generate a 227 bp fragment. These primers had the following sequences:

Set 2-5'fIX (code no. 795X): 5' GAG GAG ACA GAA CAT ACA GAG C 3'  
Set 2-3'fIX (code no. 794X): 5' CAG GTA AAA TAT GAA ATT CTC CC 3'

and were used against a variety of RNA prepared from tissues expressing fIX. The results are shown in Table 1.

TABLE 1

RNA	PCR Fragment	Splice	Comment
Human liver	689	no	normal splicing
Control m. mammary	N/A	N/A	no FIX expression
Control s. mammary	N/A	N/A	no FIX expression
BIX (FIXD + BLG)	227	yes	confirms sequence
FIXA: sheep mam	227	yes	aberrant splice also
FIXA: mouse mam	689	no	splice not observed
JFIXA1: sheep mam	227	yes	aberrant splice also

FIXA and JFIXA1 in sheep mammary gland do show the same aberrant splice as BIX, therefore it is not strictly construct dependent. FIXA in mouse does, however, present a rather confusing situation. Only 1/12 mice expressed this construct, but at relatively high levels (30 µg/ml). The mouse clearly does not carry out this aberrant splice in the mammary gland and hence quite high levels of FIX in milk are seen. But why this happens in this one mouse is not understood. Nevertheless it suggests that the absence of the aberrant splice can improve FIX levels in milk.

### EXAMPLE 3 - Construction of FIX-Δ3'splice

This construction is outlined in Figure 6. A set of PCR primers (set 4)

Set 4 5'BLG (976G) 5'GCT TCT GGG GTC TAC CAG GAA C 3'

Set 4 3'FIX (2212) 5'TAT AAC CCG GGA AAT CCA TCT TTC ATT AAG T 3'\*

\*contains additional 5' sequence including new SmaI site for cloning purposes.

were used to amplify a segment of FIXD from the 5' BLG sequence to a sequence just 3' to the stop codon of FIX but 5' to the cryptic acceptor splice site. This segment of DNA thus contains the coding sequence of FIX but lacks the cryptic acceptor site in the 3' untranslated region. This segment was fused to BLG sequences to make a construct very similar to FIXD but lacking 141 bp of 3' flanking sequences of FIX present in FIXD, including the cryptic acceptor site.

#### **EXAMPLE 4 - Expression of FIX-Δ3'splice**

To test whether FIX-Δ3'splice resulted in improved FIX expression in transgenic animals it was coinjected with BLG into mouse eggs (as per WO-A-9211385) and a number of transgenic lines established. Expression of the FIX-Δ3'splice transgene was analysed in the mammary gland at the RNA and protein level.

**Protein analysis:** Nine lines of transgenic mice have so far been analysed. All of them exhibit detectable levels of FIX in milk. One of them (line 31) showed very high levels (an average of 60.9 µg/ml) with some individuals showing >100 µg/ml (Table 2): this is by far the highest level of FIX ever achieved in milk.

#### **ELISA Analysis of Factor IX Milk Samples**

These milks were from transgenic mice with the modified factor IX cDNA (acceptor splice site removed). The ELISA is based on capture by a rabbit polyclonal and detection is by the same polyclonal but modified by biotinylation. Expression is indicated below:-

TABLE 2: RNA and Protein Expression in FIXA3' Lines

Line	Copy Nos.* BLG/FIXDA3	RNA (ng/μg)®	Protein μg/μl <sup>+</sup>	
5	3	nd	+	2.9 (2)
	11	8/2	+(.04)	4.2 (3)
	12	15/2	+(.02)	9.1 (8)
	14	14/3	-	0.3 (1)
	23	28/3	-	0.4 (2)
10	31	6/2	+(.44)	60.9 (18) \$
	34	9/1	-	0.38 (3)
	41	6/1	-	<0.1 (2)
	44	nd	+	0.6 (3)

- 15 \* estimated by PhosphorImager analysis of S. blots of tail DNA; these values are approximate ("nd" indicates "not done")
- 20 <sup>®</sup> in some samples the level of FIXDA3' mRNA was estimated relative to an *in vitro* transcribed FIX transcript
- + measured by ELISA; averaged from the number of G<sub>1</sub> (first generation) or G<sub>2</sub> (second generation) samples shown in parentheses
- 25 <sup>\$</sup> FIX levels exceeded 100  $\mu$ g/ml in some individuals of this line

Furthermore, the protein produced has a very similar mobility to normal plasma derived human FIX on reducing and non-reducing gels (Figure 7) and is biologically active (Table 3). These levels of FIX production would be commercial in sheep.

#### Purification and Biological Activity of Human FIX from Transgenic Mouse Milk

35 FIX was purified from pooled mouse milks from line 31 by immunoaffinity chromatography. MabA7 which binds the Ca<sup>2+</sup> binding FIX Gla domain was a kind gift from Charles Lutsch. The antibody was coupled to cyanogen bromide activated Sepharose. Diluted milk was incubated

overnight with antibody-conjugated Sepharose in 50 mM Tris, 150 mM NaCl pH 7.5 (TBS) + 50 mM  $\text{CaCl}_2$  at 4°C. Bound protein was eluted isocratically with TBS, 25 mM EDTA, pH 7.5 FIX coagulation activity was measured by the addition of FIX deficient plasma (Diagnostic Reagents, Oxon, UK) and APTT reagent (Sigma) with the reaction initiated after 5 minutes by addition of  $\text{Ca}^{++}$ . Coagulation was measured by ball oscillation with an ST4 Analyser (Diagnostica Stago). Normal human plasma (4  $\mu\text{g/ml}$  FIX as measured by ELISA) was used as standard. The results are indicated in Table 3 below:-

TABLE 3

Pooled Milk <sup>*</sup>	Eluate			
	Total FIX <sup>@</sup> ( $\mu\text{g}$ )	Recovery	Conc <sup>n@</sup> ( $\mu\text{g/ml}$ )	Activity <sup>+</sup> ( $\mu\text{g/ml}$ )
140	61.6	44%	30.8	30.85

\* a number of milk samples from line FIX $\Delta$ 3'31 were pooled

@ measured by ELISA

+ measured by clotting assay

RNA analysis: Northern blots of representative RNA samples from FIX- $\Delta$ 3'splice mice were probed with a FIX-specific probe. The predicted size transcripts (~1680 nt) were observed (Figure 8) and, furthermore, the steady state mRNA levels correlated with the levels of FIX detected in milk (eg line 31 had the highest mRNA levels (see Table 2)). These FIX- $\Delta$ 3'splice RNAs were co-run with some BIX RNAs. Note that they have a higher molecular weight than the BIX mRNA (1351 nt) even though the construct is smaller. The aberrant splice which shortens BIX mRNA has now been cured. This was confirmed



by an RT-PCR analysis of FIX- $\Delta 3'$  splice RNA which showed that the 3' segment of the transcript was intact (not shown).

CLAIMS

1. DNA having a sequence encoding a protein having human factor IX activity, wherein the DNA is modified to interfere with the functioning of at least one of the following cryptic splice sites:
- 5 (a) a donor site including mRNA nucleotide 1086; and
- 10 (b) an acceptor site including mRNA nucleotide 1547;
- adopting the mRNA nucleotide numbering of Figure 2 of the drawings.
2. DNA as claimed in claim 1, which encodes wild-type human factor IX.
- 15 3. DNA as claimed in claim 1 or 2, which contains at least one of the introns present in genomic DNA encoding factor IX.
- 20 4. DNA as claimed in claim 1, 2 or 3, in which the cryptic donor site is engineered out.
- 25 5. DNA as claimed in any one of claims 1 to 4, in which the cryptic acceptor site is engineered out.
- 30 6. DNA as claimed in claim 5, which is a DNA segment encoding factor IX, the DNA segment being shortened at its 3' end to exclude the acceptor site.
- 35 7. DNA as claimed in claim 6, which is cDNA.
8. An expression host comprising DNA as claimed in any one of claims 1 to 7 operably linked to an expression control sequence.

9. An expression host as claimed in claim 8, which is a transgenic non-human animal.

5 10. An expression host as claimed in claim 9, wherein the animal is a placental mammal and the expression control sequence directs expression in the mammary gland so that factor IX is present in the mammal's milk.

10 11. An expression host as claimed in claim 9, wherein the expression control sequence comprises the  $\beta$ -lactoglobulin promoter.

1/7

FIG. 1

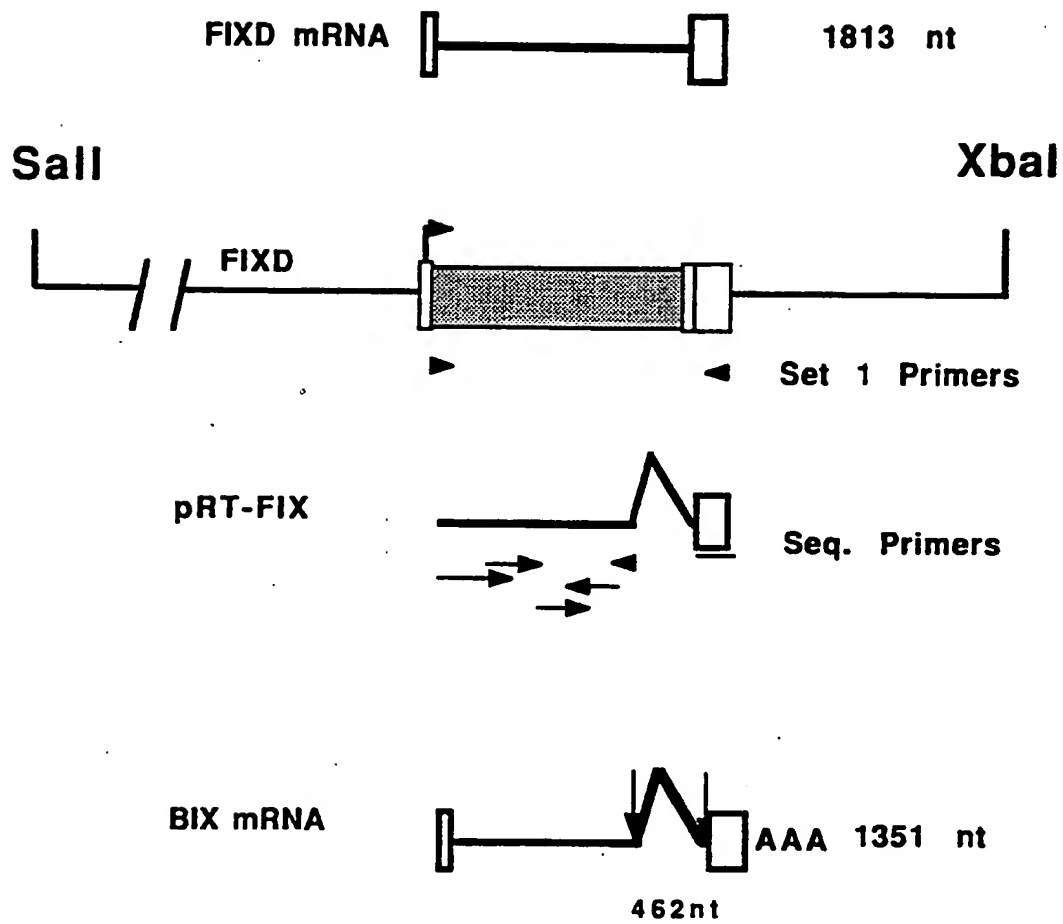




FIG. 2II

3/7

Q	Y	U	S	D	R	V	F	M	K	G	A	S	A	L	V	K	Q	Y	L	R	V	P	L	U	D	R	A	T	C	L	R	S	T	K	F	T	I
1080	1100	1110	1120	1130	1140	1150	1160	1170	1180	1190	1200	1210	1220	1230	1240	1250	1260	1270	1280	1290	1300	1310	1320	1330	1340	1350	1360	1370	1380	1390	1400	1410	1420	1430	1440	1450	1460
Y	M	N	E	C	A	D	F	N	E	O	S	R	D	S	C	O	D	S	Q	D	P	M	U	T	E	O	E	O	T	S	F	L	T	G	I	I	
1260	1270	1280	1290	1300	1310	1320	1330	1340	1350	1360	1370	1380	1390	1400	1410	1420	1430	1440	1450	1460	1470	1480	1490	1500	1510	1520	1530	1540	1550	1560	1570	1580	1590	1600	1610	1620	
N	O	E	C	A	M	K	Q	K	Y	D	I	Y	K	V	S	R	V	N	W	I	K	E	K	T	K	L	T										
1360	1370	1380	1390	1400	1410	1420	1430	1440	1450	1460	1470	1480	1490	1500	1510	1520	1530	1540	1550	1560	1570	1580	1590	1600	1610	1620	1630	1640	1650	1660	1670	1680	1690	1700	1710	1720	
G	O	G	U	P	A	D	A	G	U	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
1720	1730	1740	1750	1760	1770	1780	1790	1800	1810	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000	2010	2020	2030	2040	2050	2060	2070	2080	
A	G	U	C	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	
2080	2090	2100	2110	2120	2130	2140	2150	2160	2170	2180	2190	2200	2210	2220	2230	2240	2250	2260	2270	2280	2290	2300	2310	2320	2330	2340	2350	2360	2370	2380	2390	2400	2410	2420	2430	2440	
C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	
2440	2450	2460	2470	2480	2490	2500	2510	2520	2530	2540	2550	2560	2570	2580	2590	2600	2610	2620	2630	2640	2650	2660	2670	2680	2690	2700	2710	2720	2730	2740	2750	2760	2770	2780	2790	2800	
G	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	
2800	2810	2820	2830	2840	2850	2860	2870	2880	2890	2900	2910	2920	2930	2940	2950	2960	2970	2980	2990	3000	3010	3020	3030	3040	3050	3060	3070	3080	3090	3100	3110	3120	3130	3140	3150		
U	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	
3150	3160	3170	3180	3190	3200	3210	3220	3230	3240	3250	3260	3270	3280	3290	3300	3310	3320	3330	3340	3350	3360	3370	3380	3390	3400	3410	3420	3430	3440	3450	3460	3470	3480	3490	3500	3510	
U	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	
3510	3520	3530	3540	3550	3560	3570	3580	3590	3600	3610	3620	3630	3640	3650	3660	3670	3680	3690	3700	3710	3720	3730	3740	3750	3760	3770	3780	3790	3800	3810	3820	3830	3840	3850	3860	3870	
G	U	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	
3870	3880	3890	3900	3910	3920	3930	3940	3950	3960	3970	3980	3990	4000	4010	4020	4030	4040	4050	4060	4070	4080	4090	4100	4110	4120	4130	4140	4150	4160	4170	4180	4190	4200	4210	4220	4230	
G	U	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	
4230	4240	4250	4260	4270	4280	4290	4300	4310	4320	4330	4340	4350	4360	4370	4380	4390	4400	4410	4420	4430	4440	4450	4460	4470	4480	4490	4500	4510	4520	4530	4540	4550	4560	4570	4580	4590	
U	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	
4590	4600	4610	4620	4630	4640	4650	4660	4670	4680	4690	4700	4710	4720	4730	4740	4750	4760	4770	4780	4790	4800	4810	4820	4830	4840	4850	4860	4870	4880	4890	4900	4910	4920	4930	4940	4950	
G	U	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	
4950	4960	4970	4980	4990	5000	5010	5020	5030	5040	5050	5060	5070	5080	5090	5100	5110	5120	5130	5140	5150	5160	5170	5180	5190	5200	5210	5220	5230	5240	5250	5260	5270	5280	5290	5300	5310	
G	U	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	C	A	

SUBSTITUTE SHEET (RULE 26)

4/7

FIG. 3

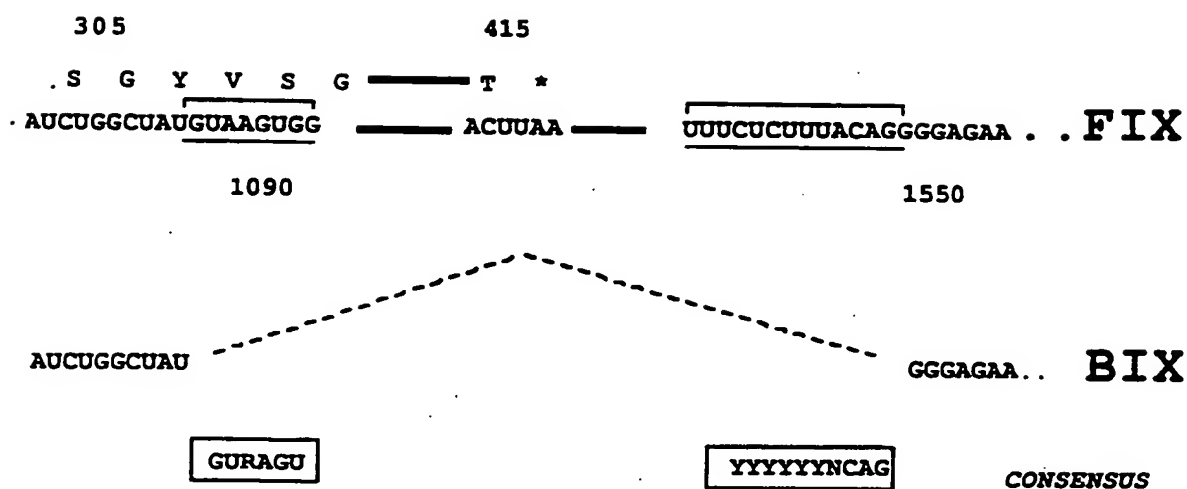
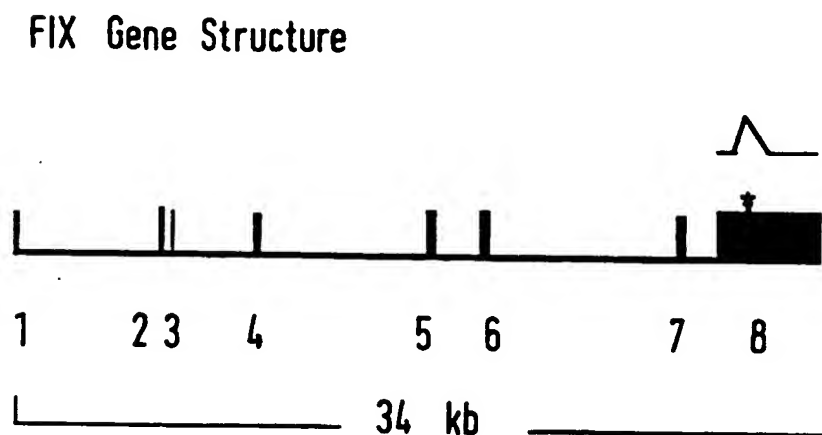
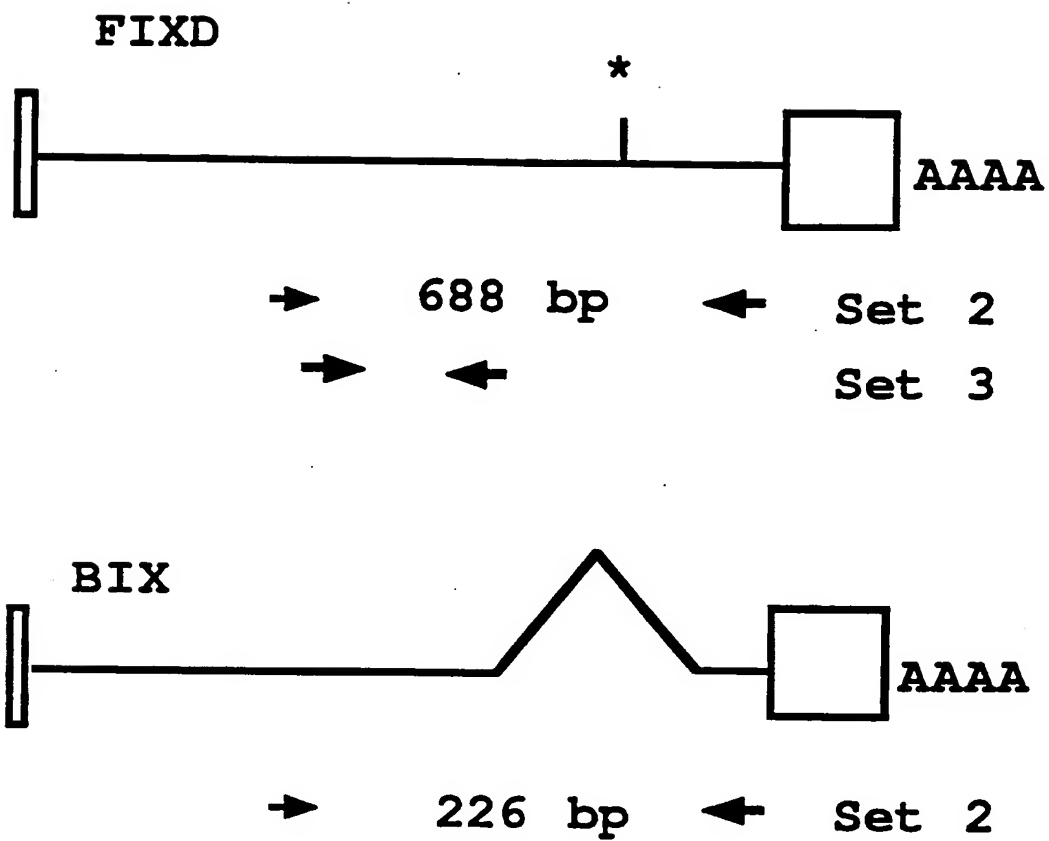


FIG. 4



5/7

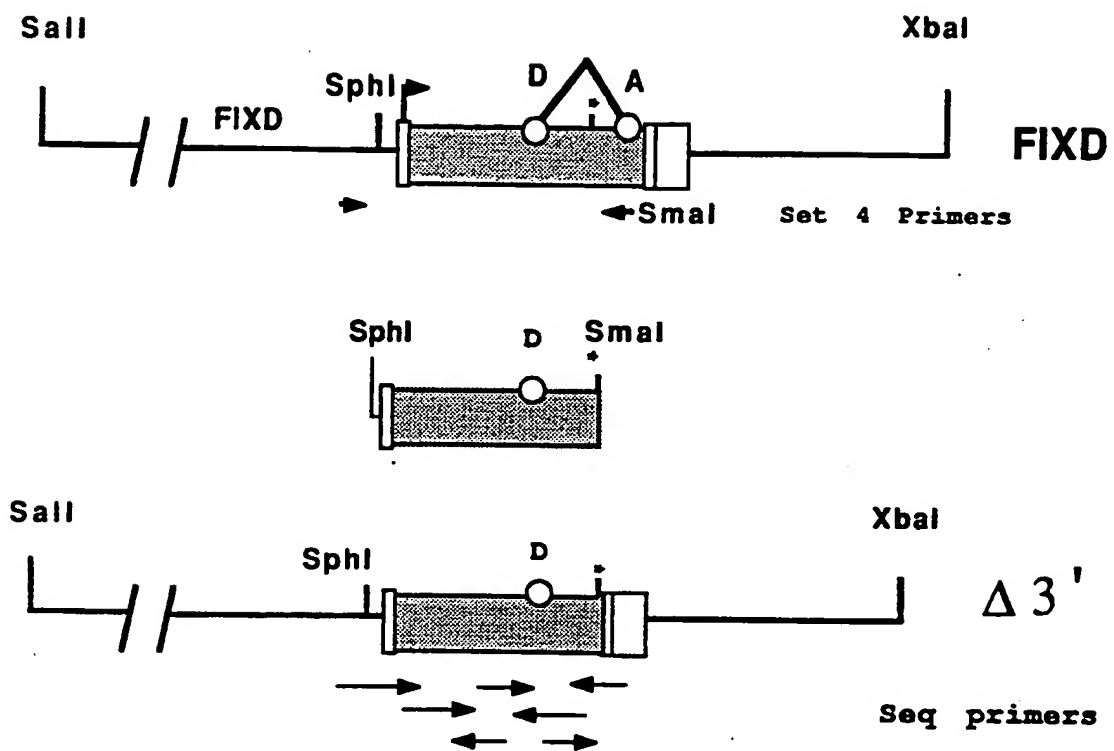
FIG. 5

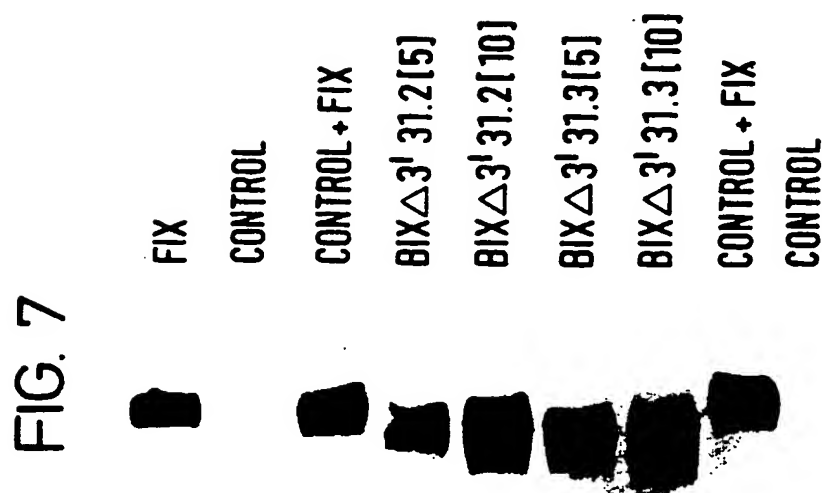
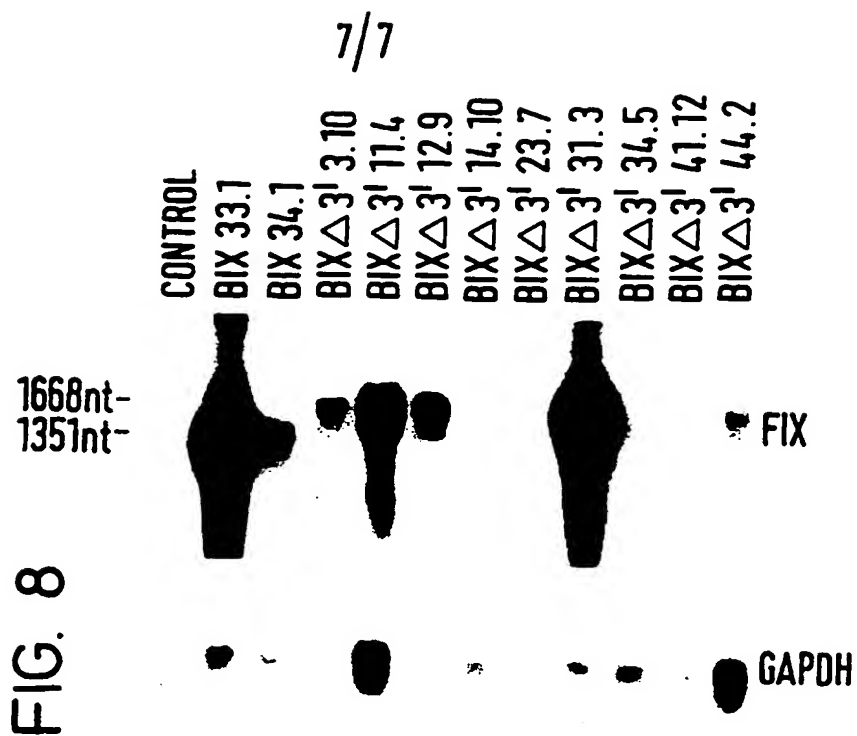




6/7

FIG. 6





# INTERNATIONAL SEARCH REPORT

Intern. Application No.  
PCT/GB 95/00996

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12N15/12 C12N15/63 C12N5/10 A01K67/027 C12N15/11

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12N A01K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	EMBO JOURNAL, vol. 3, no. 5, 1984 pages 1053-1060, D.S. ANSON ET AL. 'The gene structure of human anti-haemophilic factor IX' *see the whole article*	1-11
A	Journal of Cellular Biochemistry, 1988, Suppl. 12A, page 195, abstract H226 *see the whole abstract*	1-11
	-/--	

☒ Further documents are listed in the continuation of box C.

☐ Patent family members are listed in annex.

### \* Special categories of cited documents :

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier document but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed

- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- "&" document member of the same patent family

Date of the actual completion of the international search

4 August 1995

Date of mailing of the international search report

01.09.95

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax (+31-70) 340-3016

Authorized officer

Marie, A

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/GB 95/00996

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>AMERICAN JOURNAL OF HUMAN GENETICS, vol. 50, 1992 pages 434-437, J. SOLERA ET AL. 'Factor IX Madrid2: a deletion/insertion in Factor IX gene which abolishes the sequence of donor junction at the exonIV-intron d splice site' *see the whole article*</p> <p style="text-align: center;">---</p>	1-11
A	<p>NUCLEIC ACIDS RESEARCH, vol. 19, no. 5, 1991 page 1172 S.H. CHEN ET AL. 'Splice junction mutations in factor IX gene resulting in severe hemophilia B' *see the whole article*</p> <p style="text-align: center;">-----</p>	1-11